Evidence Of Evidence-Based Health Policy: The Politics Of Systematic Reviews In Coverage Decisions

Systematic reviews are contributing to the making of policy that conserves scarce U.S. resources.

by Daniel M. Fox

ABSTRACT: U.S. policymakers are making greater use of findings from systematic reviews, the principal product of the discipline of research synthesis. This paper summarizes the methodology and availability of systematic reviews and the brief history of their introduction to policymakers in the public and private sectors and health professionals in the United States. Then, as a case study, the paper describes how officials in a consortium of states are using systematic reviews to inform decisions about coverage for pharmaceuticals. Finally, it explores the prospects for wider use of systematic reviews by policymakers.

FINDINGS FROM SYSTEMATIC REVIEWS, the principal product of the discipline of research synthesis, are being used increasingly by U.S. policymakers. Policymakers who use systematic reviews say that such reviews help them make decisions about allocating scarce resources and responding to advocacy for covering particular drugs, procedures, and medical devices. Most of this advocacy, by disease-specific groups, manufacturers, and organizations of providers, is based on the results of one or a few studies of primary data, often accompanied by anecdotes about particular interventions' effectiveness. In contrast, a systematic review arrays, evaluates, and summarizes the aggregate results of every study of an intervention (or of competing interventions) that can be found.

This paper summarizes the methodology and availability of systematic reviews and the brief history of their introduction to policymakers in the public and private sectors and health professionals in the United States. Then, as a case study, it describes how officials in a consortium of states are using systematic reviews to inform decisions about coverage for pharmaceuticals. Finally, it explores the prospects for wider use of systematic reviews by policymakers.

Dan Fox (dmfox@milbank.org) is president of the Milbank Memorial Fund in New York City.

DOI 10.1377/hlthaff.24.1.114

Methods And Availability Of Systematic Reviews

Systematic reviews are "an old practice whose methods have been refined in recent decades within the social and health sciences," writes Ray Moynihan, a journalist who has been covering their production and use in many countries since the early 1990s.[1] Their defining characteristic is that they adhere to internationally agreed-upon standards that make them "rigorous, transparent and up-to date," according to Iain Chalmers, the British researcher who has led the international movement to establish these standards.[2] Systematic reviews identify and eliminate bias in primary studies more effectively than the less rigorous reviews of the past.

The authors of a systematic review conducted under internationally accepted standards must first devise a protocol and submit it to peers for review. The protocol establishes a specific question or questions that the review will answer. Then the authors search exhaustively for relevant research, unpublished as well as published; use rigorous evaluative criteria for identifying bias and including and excluding studies of primary data (a process called "appraisal"); and employ statistical meta-analysis, narrative synthesis, and other pertinent methods to synthesize data. Reviewers also agree to regularly update published reviews.[3]

More than 2,000 systematic reviews have been published that meet international standards. By far the most are written by members of the Cochrane Collaboration, an organization of more than 12,000 members from ninety-one countries that was founded in 1992. Four times each year, the fifty Cochrane Review Groups (CRGs) publish new and updated reviews in the electronic Cochrane Library.[4] Approximately 1,000 Cochrane reviewers are Americans. The largest North American sources of high-quality systematic reviews are the thirteen Evidence-based Practice Centers (EPCs) designated by the Agency for Healthcare Research and Quality (AHRQ).

Estimates of the cost of systematic reviews vary. Leaders of the Cochrane Collaboration estimate that the cost of an average review—mainly the time and direct expenses of reviewers and support staff—translates into roughly US$50,000. Reviews conducted by EPCs can cost $250,000 or more because they address broader questions. These estimates do not include the costs of sustaining the groups of researchers who produce and edit reviews. A recent estimate by members of the Funders Forum of the Cochrane Collaboration is that, worldwide, core funding explicitly earmarked for systematic reviewing totals around US$20 million a year, with most of it supplied—both absolute and per capita—by the governments of the United Kingdom, Australia, and the Nordic countries.[5]

Systematic Reviews And U.S. Policymakers

A few state policymakers became aware of the potential value of systematic reviews in 1990, shortly after Chalmers and his colleagues published a landmark book based on systematic reviews of an entire field of health care. *Effective Care in*

Pregnancy and Childbirth described the methodology of systematic reviews; applied it to perinatal care; and concluded with lists of interventions that were beneficial, appeared to confer no benefits, and required further research.[6] State policymakers responded enthusiastically to a description of systematic reviews and examples of beneficial and questionable interventions drawn from this book at a workshop organized by the User Liaison Program of what is now AHRQ in spring 1990.[7]

Two subsequent events suggested that interest in systematic reviews among policymakers and their constituents could grow along with the number of published reviews. In April 1991, speakers at a national conference sponsored by the American College of Obstetrics and Gynecology (ACOG), the International Society of Technology Assessment of Health Care (ISTAHC), and the Milbank Memorial Fund discussed the findings of *Effective Care in Pregnancy and Childbirth* and their implications for policy and practice. The assistant secretary for health in the U.S. Department of Health and Human Services (HHS) keynoted the conference, which attracted several hundred participants. In February 1993, systematic reviews attracted national media attention for the first time when the pioneering science journalist Earl Ubell wrote a cover story for *Parade*, the national Sunday newspaper supplement, about the book's findings.[8]

Interest in systematic reviews grew among researchers; leading medical journalists; and a few decisionmakers in government, large corporations, and health plans. By the end of the 1990s almost 1,200 reviews were available electronically in the Cochrane Library, and EPCs were receiving more commissions for reviews from government agencies and professional societies. The Blue Cross/Blue Shield Technology Assessment Project (TAP) produced reviews that some health plans used to inform coverage decisions. ECRI (originally the Emergency Care Research Institute), which became one of the EPCs, had been doing meta-analyses of studies of technology since the early 1990s that were used primarily by hospitals to inform decisions about purchasing equipment.

Since the late 1990s, medical journals have published an increasing number of systematic reviews; reporters for the *New York Times*, the *Wall Street Journal*, and the *Economist*, among others, have cited them; and the Institute of Medicine (IOM) has described the importance of systematic reviews for practice and policy.[9]

As the number of available reviews increased, evidence accumulated that they were informing decisions about coverage. Since 1999, for instance, the Centers for Medicare and Medicaid Services (CMS, formerly HCFA) has commissioned systematic reviews as a step in making national coverage decisions. Medical directors of health plans have referred to systematic reviews in describing how they made decisions about covering controversial therapies. Documented examples include coverage of pancreas transplantation and of autologous bone marrow transplant for metastatic breast cancer. Legislation in Washington State mandates that the "best available scientific and medical evidence" should guide coverage decisions for every agency of state government that purchases health care and that this evi-

dence should be "derived from systematic research."[10]

Policymakers and their allies in Wyoming greatly expanded consumers' access to systematic reviews. In 2004 the state made access to the Cochrane Library freely available to all in Wyoming. In addition, the state established a Prescription Drug Resource Center, operated in collaboration with AARP. Under this program, residents of Wyoming have online access to systematic reviews of competing drugs within classes and, for a copayment of $5, access to counseling about medications from pharmacists practicing in the state who receive special training.[11] These reviews are commissioned and published by a consortium of states called the Drug Effectiveness Review Project (DERP). The DERP offers the most compelling evidence to date of both the growing salience of systematic reviews for policy and the controversial issues raised by that salience.

The DERP And The Politics Of Policy

The DERP is a consortium that has grown to include fifteen states and two nonprofit organizations since its inception in 2001. The members of its governance body, all senior decisionmakers, establish priorities for evaluation of drug classes through systematic reviews. The project then commissions a review of each class from an EPC. Member states encourage health professionals and the general public to comment to the EPC on key questions and draft reviews. Completed reviews and periodic updates are publicly available on a Web site maintained by the Center for Evidence-based Policy of the Oregon Health and Science University, which manages the project on behalf of the consortium.[12] John Kitzhaber, former governor of Oregon, directs the center.

■ **Use of review findings.** DERP members use review findings in different ways. Most of the states use the reviews to guide policy for Medicaid preferred drug lists. Several states use them to inform decisions about public employees' coverage. At least one state is using them to inform workers' compensation coverage decisions. Groups of physicians in North Carolina, organized regionally, recommend action to their peers based on the reviews. Wyoming seems to be the only state that encourages consumers to use DERP reviews. Drug coverage policymakers in Canadian provinces are members of the DERP through the Canadian Coordinating Office for Healthcare Technology Assessment (CCOHTA). In several provinces, such as Saskatchewan, committees of pharmacists and physicians are using the DERP reviews to inform their advice to health ministries about prescription drug coverage.

■ **Range of coverage decisions.** Because the DERP reports the findings of each of the systematic reviews it commissions in probabilistic language and does not make recommendations for policy, each review can lead to different coverage decisions. For example, according to Mark Gibson, who coordinates the project for the Oregon-based center, "Early on Oregon and Washington used the same review of Triptans to reach different conclusions." Because Oregon accorded priority to the "outcome of patients being 'pain free at two hours,' its policymakers chose

rizatriptan," he said. Policymakers in Washington State, he continued, "chose both rizatriptan and sumatriptan after considering a broader array of clinical endpoints."[13]

Similarly, John Santa, medical director of the center, recalled that most members decided to "go for the best price" in choosing among proton pump inhibitors (PPIs) because the review found "all the competing PPIs to be equivalent." Santa also noted that choosing which drug in a class to cover is further complicated because suppliers offer states "different deals," so that various drugs are chosen.[14]

■ **Status of DERP reviews.** By September 2004 the DERP had completed twelve reviews and updates of most of them, had ten reviews in various stages of preparation, and would soon select three additional classes for review. The project sets priorities for reviews mainly according to how much members are spending for particular classes of drugs. Available reviews assess angiotensin-converting enzyme (ACE) inhibitors, angiotensin II receptor antagonists (AIIRAs), antidepressants (second generation), beta adrenergic blockers, calcium-channel blockers, estrogens, inhaled corticosteroids, opioids (long acting), nonsteroidal anti-inflammatory drugs (NSAIDs), oral hypoglycemics, PPIs, skeletal muscle relaxants, statins, triptans, and urinary incontinence drugs.

■ **International collaboration.** The DERP is a result of the convergence of the politics of health policy in a few U.S. states and Canada with the international movement to produce systematic reviews that meet rigorous standards. In 2000 the Cochrane Collaboration and the Milbank Memorial Fund commissioned case studies from six countries on the use of systematic reviews, and population-based research more generally, to inform policy.[15] One case described the history of policy for an evidence-based preferred drug list in British Columbia.

Late in 2000 Oregon policymakers sought to close a projected $60 million shortfall in the Medicaid budget for 2001. Santa, then a state official, asked colleagues from other western states what they knew about the British Columbia list. A staff member of the Milbank Fund who attended this meeting sent him a draft of the study. A month later, Governor Kitzhaber's staff invited the B.C. policymaker who coauthored the case study to meet with them in Salem.

This meeting led to the drafting of a bill to permit the executive branch to create a drug formulary for Medicaid based on evidence of effectiveness. The executive branch decided to commission systematic reviews to guide its decisions. The state asked an existing state advisory commission, mainly comprising physicians, to interpret the evidence in each review and make recommendations to the state's Medicaid agency about the relative effectiveness of drugs in each class. Commission members would not have access to data about the cost to the state of each drug, to avoid the accusation that, formally or informally, they were influenced by cost-effectiveness criteria.

The bill encountered opposition in the legislature mainly as a result of lobbying by and on behalf of pharmaceutical companies. In late July, during the final week

of the legislative session, Kitzhaber intervened to encourage passage of the bill. According to Pam Curtis, assistant director of the center, who was then on the governor's staff, Kitzhaber told legislative leaders that he was prepared to veto the state's entire human services budget and then conduct a public campaign on behalf of the bill, after which he would convene a special session of the legislature.[16] The bill passed.

A few months later policymakers from Idaho and Washington State offered to collaborate with Oregon in commissioning systematic reviews. These states augmented funds appropriated by Oregon. Early in 2003 DERP leaders from Idaho, Oregon, and Washington State invited colleagues from other areas to join them in governing and financing the project. The founding states and their initial collaborators agreed to prorate each member's share of the project's cost for three years according to the number of jurisdictions that participated. By November 2004 Alaska, Arkansas, Kansas, Michigan, Minnesota, Missouri, North Carolina, Wisconsin, and Wyoming had joined. Several other states were negotiating contracts with the center or seeking legislation enabling them to do so. The California HealthCare Foundation (CHCF) joined on behalf of itself and the California Public Employees Retirement System (CalPERS). The CCOHTA joined on behalf of the Canadian provinces and its federal government.

■ **The DERP and the drug industry.** Relationships between the DERP and the drug industry have been uneasy. From its inception, the DERP invited drug manufacturers to volunteer results of unpublished clinical trials to the EPCs that produced and updated the reviews. The center has responded to every manufacturer's comment on published reviews. Nevertheless, industry lobbyists have encouraged action in several states to constrain the project. For instance, legislators in several states reported that lobbyists told them that the project's reviews led to unnecessary deaths and that, as Santa recalls, "States will be sued for negligence if a patient dies because of any of these decisions."[17] Other company representatives argued that if there is no evidence of major differences among competing drugs, state policy should not be to purchase the drug that has the lowest price.

Lobbying by the pharmaceutical industry has had mixed results. A bill to end Idaho's membership in the DERP failed after a legislator demanded that each witness at a public hearing on the bill disclose the sources of his or her income and the percentage of the total received from each source.[18] On the other hand, in August 2004 advocates for disease-specific groups and the pharmaceutical industry persuaded New York legislative leaders to delete a budget item that would have created a preferred drug list and membership in the DERP.

Recent events suggest that the industry acknowledges, for now, the DERP's scientific legitimacy. This forbearance may be related to accumulating evidence that systematic reviews of drug classes, when translated into coverage policy, can change market share and increase supplemental rebates to the states from drug makers.[19] For instance, an executive branch official in one state commented that

since the DERP began to publish reviews, "the companies send scientists instead of lobbyists to persuade me to cover their drugs." The commissioner of human services in another state estimates that states that use preferred drug lists receive supplemental rebates one-third larger than those in states that do not.

DERP members and center staff met with industry representatives in June 2004 to discuss, as Gibson describes it, how to improve communication "from a process and scientific focus." The DERP agreed to disclose drafts of key questions to be addressed by reviews one month before they become final and, Gibson said, "give the interests and companies a chance to comment" on them. Moreover, he continued, the DERP will release drafts of reviews two weeks after "our participants have seen them and anyone can comment on them as long as they use the format" established by the project and the participating EPCs. At the end of July 2004, the project posted draft reviews on its Web site and invited public comment.[20]

The Future Of Systematic Reviews In Health Policy

Systematic reviews are contributing to U.S. decision making, but there are impediments to the expansion of their use: lack of understanding of the methodology and value of reviews among health professionals, the general public, and most policymakers; limitations on the supply of reviews as a result of inadequate core funding to produce them; and the negative effects of exaggeration by some systematic reviewers and other "evidence-based" researchers of the superiority of rigorous statistical evidence about populations to other forms of evidence.

In the absence of widespread agreement about the usefulness of systematic reviews, advocates for people with serious chronic diseases, the drug companies that often subsidize them, and the health professionals who endorse their claims are likely to continue their success in lobbying for access to treatments that may not be justified by the best available evidence.[21] In a number of DERP member states, these groups succeeded in excluding specific classes of drugs from preferred drug lists.

■ **Importance of media coverage.** The future of systematic reviews as a tool for making policy depends a great deal on how effectively both medical journals and journalists in general print and electronic media communicate about them. Favorable media response to proposals for a national registry of clinical trials by the American Medical Association, and endorsement of this proposal by editors of leading medical journals, may lead to growing public understanding of the importance of access to the best available evidence, both published and unpublished. Some media coverage of the proposed trials registry has linked it explicitly to systematic reviews.[22] Another promising development is the recent establishment of a news service about systematic reviews by the Center for the Advancement of Health.

A related impediment to public and professional understanding is that many of the most reliable systematic reviews of health care technology produced in the United States are neither in the public domain nor available to public and medical

libraries and individuals by subscription, as the Cochrane Library is. Credible nonprofit organizations, like ECRI and the Blue Cross/Blue Shield TAP, must charge high access fees to subscribers to recover the costs of reviews.

■ **Need for more trained researchers.** An increased supply of timely and reliable reviews and updates of them requires more people who have been trained in the methodology of research synthesis. About 500 reviews that meet international standards of quality are now published and accessible each year. This number is considerably below potential demand. Members of the international Steering Group of the Cochrane Collaboration estimated in 2003 that about 10,000 reviews would be required to assess the current array of health care interventions. Unanticipated advances in laboratory and clinical research are likely to increase that estimate.

■ **Research funding priorities.** The supply of new systematic reviews and timely updates is governed by the funding policies of the federal government and major foundations. Funding has been increasing both to commission systematic reviews and to provide core support for groups that produce them in the United States. But systematic reviewing is not yet a priority of research policy, especially at the National Institutes of Health. As the IOM Clinical Research Roundtable concluded, the "science of research synthesis, which provides the core of evidence-based medicine, ...has lagged behind that of other nations."[23] This lag seems to be a result of competing priorities for research funding and the relatively recent development of the methodology of research synthesis rather than of any organized opposition to the production of systematic reviews.

■ **Narrow versus broader definition of evidence.** Another impediment may be the excessive zeal of some proponents of evidence-based medicine. More than a few clinicians and health services researchers complain that some advocates of "evidence-based" research and practice use a narrow definition of *evidence*. They insist that policy and practice must be based on a broader array of evidence than statistical inference about events in populations that are studied prospectively. They emphasize the contributions to policy and practice of evidence from observational studies, the analysis of administrative data, and a variety of qualitative methods.

Fortunately, this critique is endorsed by many members of the Cochrane Collaboration and researchers affiliated with EPCs. Groups of researchers from many countries are devising methods to expand the types of evidence, including evidence from observational and qualitative studies, used in systematic reviews while maintaining their persuasiveness.[24]

■ **Antagonism from providers.** Finally, systematic reviews that question the effectiveness of particular interventions provoke antagonism from physicians and hospitals that are reimbursed to provide them. A cautionary example was the backlash in 2001 against a review of mammography that was controversial among reviewers as well as physicians and advocates for women's health.[25]

In sum, systematic reviews are contributing to the making of policy in the United States that conserves scarce resources. But election and news cycles may

frequently be too short to permit the best science to inform the politics of policy making. Future reports on the evidence about evidence-based policy making are likely to include examples of frustration as well as of success.

NOTES

1. R. Moynihan, *Evaluating Health Services: A Reporter Covers the Science of Research Synthesis* (New York: Milbank Memorial Fund, 2004), 2.
2. I. Chalmers, "Trying to Do More Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, and Up-to-Date Evaluations," *Annals of the American Academy of Political and Social Science* 589, no. 3 (2003): 37.
3. Moynihan, *Evaluating Health Services*, 14; and A. Boaz, D. Ashby, and K. Young, "Systematic Reviews: What Have They Got to Offer Evidence based Policy and Practice?" Working Paper no. 2 (London: ESCRC U.K. Centre for Evidence-based Policy and Practice, January 2002).
4. Cochrane Library, www.thecochranelibrary.com (effective 2 January 2005).
5. Data from internal documents of the Cochrane Funders Forum. The author is co-convener of the forum.
6. I. Chalmers, M. Enkin, and M. Keirse, *Effective Care in Pregnancy and Childbirth*, 2 vols. (Oxford: Oxford University Press, 1989).
7. The author was the presenter.
8. E. Ubell, "Are Births as Safe as They Could Be?" *Parade*, 7 February 1993.
9. Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the Twenty-first Century* (Washington: National Academies Press, 2001), 147; and N. Sung et al., "Central Challenges Facing the National Clinical Research Enterprise," *Journal of the American Medical Association* 289, no. 10 (2003): 1278–1287.
10. Washington State Senate Bill 6088, Chapter 29, Laws of 2003; House Bill 1299, Chapter 276, Laws of 2003, signed by Governor Gary Locke, effective 27 July 2003 (Rep. Eileen Cody provided this citation); and Washington Administrative Code for Pharmacy, WAC 388-530-1050 (Siri Charles of the Washington Department of Social and Health Services provided this citation).
11. AARP, Prescription Drug Resource Center, www.aarp.org/wy/rx (13 October 2004).
12. OHSU, Drug Effectiveness Review Center, www.ohsu.edu/drugeffectiveness (13 October 2004).
13. Mark Gibson, Center for Evidence-based Policy, e-mail message and interviews, 11 and 12 July 2004.
14. John Santa, Center for Evidence-based Policy, e-mail message, 12 July 2004.
15. D.M. Fox and A.D. Oxman, eds., *Informing Judgment: Case Studies of Health Policy and Research in Six Countries* (New York: Milbank Memorial Fund and Cochrane Collaboration, 2001).
16. Pam Curtis, Center for Evidence-based Policy, telephone interview, 18 July 2004.
17. Santa, e-mail message, 12 July 2004. Also, undated PowerPoint presentation by J. Garner, "Oregon's 'Evidence'-based Drug Evaluation Process: Fact, Fiction, and Liability Exposure."
18. Margaret Henbest, Idaho state representative, interview, 1 June 2004.
19. Medical Assistance Administration, Washington State, *Percent Preferred Drug Market Share*, 1 June 2004. This internal report, prepared by Siri Charles, arrays market share for eight classes of drugs between September 2003 and May 2004.
20. Gibson, e-mail message, 12 July 2004; and Curtis, telephone interview.
21. T. Craig, "Community Leaders Decry Lobby Firm's Fax," *Baltimore Sun*, 9 March 2002.
22. B. Meier, "AMA Adds Its Voice to Call for Disclosure on Drug Trials," *New York Times*, 16 June 2004.
23. Sung et al., "Central Challenges," 1284.
24. See *Cochrane Collaboration Methods Group Newsletter*, vol. 8, June 2004 (and previous annual issues); and Canadian Health Services Research Foundation, "The Science of Research Synthesis: Making It Useful for Evidence-based Management and Policy-Making," Request for Proposals (Ottawa: June 2003).
25. O. Olsen and P. Goetsche, "Screening for Breast Cancer with Mammography," *Cochrane Library*, no. 4 (2001). See also M. Petticrew, "Why Certain Systematic Reviews Reach Uncertain Conclusions," *British Medical Journal* 326, no. 7392 (2003): 756–758.